Available online at www.sciencedirect.com

SCIENCE DIRECT®

LINEAR ALGEBRA AND ITS APPLICATIONS

ELSEVIER

Linear Algebra and its Applications 374 (2003) 219–230

www.elsevier.com/locate/laa

Low rank perturbations and the spectrum of a tridiagonal sign pattern

L. Elsner [a], D.D. Olesky [b,*,1], P. van den Driessche [c,1]

[a] *Fakultät für Mathematik, Universität Bielefeld, Postfach 100131, 33501 Bielefeld, Germany*

[b] *Department of Computer Science, University of Victoria, P.O. Box 3055, Victoria, BC, Canada V8W 3P6*

[c] *Department of Mathematics and Statistics, University of Victoria, Victoria, BC, Canada V8W 3P4*

Received 11 August 2002; accepted 6 May 2003

Submitted by D. Hershkowitz

Abstract

The n-by-n tridiagonal sign pattern T_n has every superdiagonal entry positive, every subdiagonal entry negative, the $(1, 1)$ entry negative, the (n, n) entry positive and every other diagonal entry zero. Inertia and spectral results for matrices A_n having the sign pattern T_n are proved using new techniques on low rank perturbations. It is also shown (by using MAPLE) that for $8 \leqslant n \leqslant 16$, T_n allows any spectrum. These results extend those previously in the literature, and strengthen the conjecture that T_n allows any spectrum for all values of n. In addition, bounds on the algebraic multiplicity of an eigenvalue of a low rank perturbation of a general matrix are obtained.

AMS classification: 15A18

Keywords: Algebraic multiplicity; Inertia; Low rank perturbation; Sign pattern; Spectrum; Tridiagonal matrix

* Corresponding author. Tel.: +1-250-721-7223; fax: +1-250-721-7292.

E-mail addresses: elsner@mathematik.uni-bielefeld.de (L. Elsner), dolesky@cs.uvic.ca (D.D. Olesky), pvdd@math.uvic.ca (P. van den Driessche).

[1] Research supported in part by an NSERC Research grant.

doi:10.1016/S0024-3795(03)00580-9

1. Introduction

In [2] the n-by-n antipodal tridiagonal sign pattern T_n with $n \geqslant 2$ was introduced, where

$$T_n = \begin{bmatrix} - & + & 0 & \cdots & 0 \\ - & 0 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & + \\ 0 & \cdots & 0 & - & + \end{bmatrix}.$$

For inertia and spectral results it suffices to represent a real matrix A_n having the sign pattern of T_n by $n+1$ positive real numbers $a_0, \ldots, a_n$, but for some purposes it is more convenient to stress the skew symmetry of the nondiagonal entries. Thus we write

$$A_n = \begin{bmatrix} -a_0 & 1 & 0 & \cdots & 0 \\ -a_1 & 0 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & 1 \\ 0 & \cdots & 0 & -a_{n-1} & a_n \end{bmatrix}$$

or (1)

$$A_n = \begin{bmatrix} -\alpha_0^2 & \alpha_1 & 0 & \cdots & 0 \\ -\alpha_1 & 0 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & \alpha_{n-1} \\ 0 & \cdots & 0 & -\alpha_{n-1} & \alpha_n^2 \end{bmatrix}$$

in which $a_j > 0$, $\alpha_j > 0$ and $\alpha_j^2 = a_j$ for all j. Sometimes it is convenient to write these as $A_n(a_0, \ldots, a_n)$ or $A_n(\alpha_0, \ldots, \alpha_n)$, respectively. The *inertia* of T_n is the set of all possible ordered triples $i(A_n) = (i_+(A_n), i_-(A_n), i_0(A_n))$ giving the number of eigenvalues of $A_n \in \mathbb{R}^{n \times n}$ with positive, negative, zero real parts, respectively.

In [2] some inertia and spectral results for T_n are proved; in particular, for $2 \leqslant n \leqslant 7$ it is shown that T_n allows any spectrum (and thus any inertia), and this is conjectured to hold for all n. This problem can be considered as a particular inverse eigenvalue problem; for an overview of inverse eigenvalue problems see [1].

Our aim is to extend the results in [2] by developing new techniques based mainly on characteristic polynomials and low rank perturbations. We address some of the questions posed in [2] and our answers strengthen the above conjecture. Recently some patterns that allow any spectrum [7] or any inertia [3] have been displayed.

However, we remark that A_n depends only on $n+1$ positive parameters, none of which can be set to zero without making the resulting matrix reducible or have eigenvalues strictly in one half plane. Finally, low rank perturbations of general matrices are considered. In particular, bounds on the algebraic multiplicity of an eigenvalue in terms of the rank of the perturbation are given.

2. Inertia results for T_n

If $a_j = a_{n-j}$ or $\alpha_j = \alpha_{n-j}$ for $j = 0, 1, \ldots, n$, then we say that A_n is *symmetric about the reverse diagonal*. Define the n-by-n matrix

$$S_n = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ -a_1 & 0 & \ddots & & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & 0 & 1 \\ 0 & \cdots & 0 & -a_{n-1} & 0 \end{bmatrix},$$

and the polynomial

$$s_{ik}(\lambda) = \det(\lambda I_{k-i+1} - S_n[i, \ldots, k]), \tag{2}$$

where $S_n[i, \ldots, k]$ denotes the principal submatrix from rows and columns i to k $(k \geqslant i)$ of S_n and I_{k-i+1} denotes the identity matrix of order $k-i+1$. Thus s_{1n} is the characteristic polynomial of S_n. Since S_n is diagonally similar to a skew symmetric matrix, s_{1n} is an odd or even polynomial in λ according as n is odd or even. With the above notation the following equivalences hold.

Theorem 2.1. *For $A_n(a_0, \ldots, a_n)$ of the form* (1), *the following are equivalent*:

(i) $\det(\lambda I_n - A_n)$ *is an odd or even polynomial in λ according as n is odd or even.*
(ii) $a_0 s_{2n}(\lambda) = a_n s_{1,n-1}(\lambda)$, *with $s_{ik}(\lambda)$ given by* (2).
(iii) A_n *is symmetric about the reverse diagonal.*

Proof. The equivalence of (i) and (ii) follows from the determinantal expansion

$$\det(\lambda I_n - A_n) = (s_{1n}(\lambda) - a_0 a_n s_{2,n-1}(\lambda)) + (a_0 s_{2n}(\lambda) - a_n s_{1,n-1}(\lambda))$$

in which the first parenthesized term is an odd or even polynomial in λ according as n is odd or even, and the second parenthesized term is even or odd according as n is odd or even.

To show the equivalence of (ii) and (iii), consider the statements

$$C(k) : a_i = a_{n-i} \text{ for } i < k, \quad \text{and} \quad s_{k,n-k}(\lambda) = s_{k+1,n-k+1}(\lambda).$$

Clearly $C(1)$ is equivalent to (ii). If $n = 2m$, then $C(m)$ gives $a_i = a_{n-i}$ for $i < n/2$ and $s_{m,m}(\lambda) = s_{m+1,m+1}(\lambda) = \lambda$, thus $C(m)$ is equivalent to (iii). Similarly if $n =$

$2m+1$, then $C(m)$ gives $a_i = a_{n-i}$ for $i \leqslant m-1$ and $s_{m,m+1}(\lambda) = s_{m+1,m+2}(\lambda)$. This last equation gives $\lambda^2 + a_m = \lambda^2 + a_{m+1}$, thus $a_m = a_{m+1}$, and hence $C(m)$ is again equivalent to (iii). These latter two equivalences give (iii) is equivalent to $C(m)$ for $m = \lfloor \frac{n}{2} \rfloor$. To complete the proof of the equivalence of (ii) and (iii), we now show that $C(k)$ is equivalent to $C(k+1)$ for $1 \leqslant k \leqslant \lfloor \frac{n}{2} \rfloor - 1$. Assuming $C(k)$ gives $a_i = a_{n-i}$ for $i < k$ and $s_{k,n-k}(\lambda) = s_{k+1,n-k+1}(\lambda)$, which on expansion gives

$$\lambda s_{k+1,n-k}(\lambda) + a_k s_{k+2,n-k}(\lambda) = \lambda s_{k+1,n-k}(\lambda) + a_{n-k} s_{k+1,n-k-1}(\lambda).$$

The last equality is equivalent to $a_k = a_{n-k}$ and $s_{k+1,n-k-1}(\lambda) = s_{k+2,n-k}(\lambda)$, thus giving $C(k+1)$. The fact that $C(k+1)$ implies $C(k)$ follows by reversing the argument, thus completing the proof. $\square$

Theorem 2.1 gives a simpler proof than [2, Theorem 11] of the fact that if A_n is nilpotent (i.e., $\det(\lambda I_n - A_n) = \lambda^n$), then it is symmetric about the reverse diagonal. Theorem 2.1 is also used in the proof of Theorem 2.2, in which the construction is simpler and slightly different from that suggested in [2, Conjecture 8] but can be adapted to prove this conjecture.

Theorem 2.2. *Given $A_n(\alpha_0, \ldots, \alpha_n)$ in the form (1), fix $k \in \{1, 2, \ldots, \lfloor \frac{n}{2} \rfloor\}$. If $\alpha_j = 1$ for all j except $\alpha_k = \alpha_{n-k} = \varepsilon$, then $i(A_n) = (k, k, n-2k)$ for all sufficiently small $\varepsilon > 0$.*

Proof. For $\varepsilon = 0$, A_n is a direct sum of three matrices

$$B_k = \begin{bmatrix} -1 & 1 & 0 & \cdots & \cdots & 0 \\ -1 & 0 & 1 & \ddots & & \vdots \\ 0 & -1 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & -1 & 0 & 1 \\ 0 & \cdots & \cdots & 0 & -1 & 0 \end{bmatrix},$$

$$C_k = \begin{bmatrix} 0 & 1 & 0 & \cdots & \cdots & 0 \\ -1 & 0 & 1 & \ddots & & \vdots \\ 0 & -1 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & \ddots & -1 & 0 & 1 \\ 0 & \cdots & \cdots & 0 & -1 & 1 \end{bmatrix},$$

$$D_{n-2k} = \begin{bmatrix} 0 & 1 & & & 0 \\ -1 & 0 & 1 & & \\ & -1 & \ddots & \ddots & \\ & & \ddots & \ddots & 1 \\ 0 & & & -1 & 0 \end{bmatrix},$$

where $B_1 = [-1]$, $C_1 = [1]$ and D_{n-2k} is vacuous (i.e., A_n is a direct sum of two matrices) if $n = 2k$. For $k \geqslant 2$, matrices B_k and C_k are Schwarz matrices, and it is well known that $i(B_k) = (0, k, 0)$ and $i(C_k) = (k, 0, 0)$ [2,4]. Also $i(D_{n-2k}) = (0, 0, n-2k)$ and the eigenvalues of D_{n-2k} are distinct. For $\varepsilon > 0$, since A_n is symmetric about the reverse diagonal it follows from Theorem 2.1 that $\det(\lambda I_n - A_n)$ is an odd or even polynomial in λ according as n is odd or even. Thus if λ is an eigenvalue of A_n, so is $-\lambda$. Consequently, by continuity, for sufficiently small $\varepsilon > 0$, $i(A_n)$ must be equal to the inertia when $\varepsilon = 0$, namely $(k, k, n-2k)$. □

We now consider low rank perturbations that enable us to construct from A_n matrices of higher dimension that have purely imaginary eigenvalues. In the following theorem, matrix M with $\operatorname{rank}(M) = 1$ is used k times to give a perturbation of rank k on a block diagonal matrix consisting of $k+1$ blocks of A_n, as in (3) below. Analogous to $s_{ik}(\lambda)$, we define $p_{ik}(\lambda) = \det(\lambda I_{k-i+1} - A_n[i, \ldots, k])$. Let O_n denote the n-by-n zero matrix, o_n denote the n-by-1 zero vector, and e_j denote the n-by-1 unit vector with all entries 0 except that the jth entry is 1. The adjoint of a nonsingular matrix X is denoted by $\operatorname{adj}(X)$, i.e., $X \operatorname{adj}(X) = \det(X)I$; $[\operatorname{adj}(X)]_{ij}$ denotes its (i, j) entry.

Theorem 2.3. *For $k \geqslant 1$ and A_n of the form (1), let $C = \operatorname{diag}(A_n, \ldots, A_n) \in \mathbb{R}^{(k+1)n,(k+1)n}$. Construct matrix B with the sign pattern $T_{(k+1)n}$ by*

$$B = C + \operatorname{diag}(O_{n-1}, M, O_{n-2}, M, O_{n-2}, \ldots, M, O_{n-1}) \tag{3}$$

with $M = [\alpha_n \ \ \alpha_0]^{\mathrm{T}}[-\alpha_n \ \ \alpha_0] \in \mathbb{R}^{2\times 2}$ appearing k times in the block diagonal matrix. Then the spectrum of B contains the spectrum of A_n and all other nk eigenvalues are purely imaginary.

Proof. Let

$$\begin{aligned} w_1 &= \alpha_n(\lambda I_n - A_n)^{-1}e_n, \\ w_2 &= \alpha_0(\lambda I_n - A_n)^{-1}e_1, \\ u_i &= \left(o_{(i-1)n}^{\mathrm{T}}, \alpha_n e_n^{\mathrm{T}}, \alpha_0 e_1^{\mathrm{T}}, o_{(k-i)n}^{\mathrm{T}}\right)^{\mathrm{T}} \quad \text{for } 1 \leqslant i \leqslant k, \\ v_i &= \left(o_{(i-1)n}^{\mathrm{T}}, -\alpha_n e_n^{\mathrm{T}}, \alpha_0 e_1^{\mathrm{T}}, o_{(k-i)n}^{\mathrm{T}}\right)^{\mathrm{T}} \quad \text{for } 1 \leqslant i \leqslant k, \end{aligned}$$

and

$$\overline{u}_i = (\lambda I_{(k+1)n} - C)^{-1}u_i = \left(o_{(i-1)n}^{\mathrm{T}}, w_1^{\mathrm{T}}, w_2^{\mathrm{T}}, o_{(k-i)n}^{\mathrm{T}}\right)^{\mathrm{T}} \quad \text{for } 1 \leqslant i \leqslant k.$$

Then

$$\sum_{i=1}^{k} u_i v_i^{\mathrm{T}} = \mathrm{diag}(O_{n-1}, M, O_{n-2}, M, O_{n-2}, \ldots, M, O_{n-1})$$

and

$$\begin{aligned}
\alpha &:= v_i^{\mathrm{T}} \overline{u}_i = -\alpha_n^2 e_n^{\mathrm{T}} (\lambda I_n - A_n)^{-1} e_n + \alpha_0^2 e_1^{\mathrm{T}} (\lambda I_n - A_n)^{-1} e_1 \quad \text{for } 1 \leqslant i \leqslant k,\\
\beta &:= v_i^{\mathrm{T}} \overline{u}_{i+1} = \alpha_0 \alpha_n e_1^{\mathrm{T}} (\lambda I_n - A_n)^{-1} e_n \quad \text{for } 1 \leqslant i \leqslant k-1,\\
\gamma &:= v_i^{\mathrm{T}} \overline{u}_{i-1} = -\alpha_0 \alpha_n e_n^{\mathrm{T}} (\lambda I_n - A_n)^{-1} e_1 \quad \text{for } 2 \leqslant i \leqslant k,
\end{aligned}$$

and

$$v_i^{\mathrm{T}} \overline{u}_j = 0 \quad \text{for } |i-j| > 1.$$

Now define

$$\begin{aligned}
p(\lambda) &= \det(\lambda I_n - A_n),\\
q(\lambda) &= \det(\lambda I_{(k+1)n} - B),\\
r(\lambda) &= q(\lambda)/p(\lambda),
\end{aligned}$$

and two matrices of order $(k+1)$ n-by-k

$$\overline{U} = [\overline{u}_1, \overline{u}_2, \ldots, \overline{u}_k], \qquad V = [v_1, v_2, \ldots, v_k].$$

It follows from (3) that

$$\begin{aligned}
q(\lambda) &= \det\left(\lambda I_{(k+1)n} - C - \sum_{i=1}^{k} u_i v_i^{\mathrm{T}}\right)\\
&= \det(\lambda I_{(k+1)n} - C) \det\left(I_{(k+1)n} - \sum_{i=1}^{k} \overline{u}_i v_i^{\mathrm{T}}\right)\\
&= (p(\lambda))^{k+1} \det\left(I_{(k+1)n} - \overline{U} V^{\mathrm{T}}\right)\\
&= (p(\lambda))^{k+1} \det\left(I_k - V^{\mathrm{T}} \overline{U}\right),
\end{aligned}$$

since the nonzero eigenvalues of $\overline{U} V^{\mathrm{T}}$ and $V^{\mathrm{T}} \overline{U}$ are the same. Using the notation above, it can be seen that

$$I_k - V^{\mathrm{T}} \overline{U} = \begin{bmatrix}
1-\alpha & -\beta & & & \\
-\gamma & 1-\alpha & -\beta & & 0\\
 & \ddots & \ddots & \ddots & \\
 & & \ddots & \ddots & -\beta\\
0 & & & -\gamma & 1-\alpha
\end{bmatrix},$$

and thus

$$r(\lambda) = (p(\lambda))^k \det\left(I_k - V^{\mathrm{T}}\overline{U}\right)$$
$$= \det \begin{bmatrix} (1-\alpha)p(\lambda) & -\beta p(\lambda) & & 0 & \\ -\gamma p(\lambda) & (1-\alpha)p(\lambda) & -\beta p(\lambda) & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & -\beta p(\lambda) \\ 0 & & & -\gamma p(\lambda) & (1-\alpha)p(\lambda) \end{bmatrix}.$$

It follows from the definitions above and the structure of A_n that

$$\begin{aligned} -\gamma p(\lambda) &= \alpha_0\alpha_n e_n^{\mathrm{T}}(\lambda I_n - A_n)^{-1} e_1 p(\lambda) \\ &= \alpha_0\alpha_n[\operatorname{adj}(\lambda I_n - A_n)]_{n1} = \alpha_0\alpha_n\alpha_1\alpha_2\cdots\alpha_{n-1}, \\ -\beta p(\lambda) &= -\alpha_0\alpha_n e_1^{\mathrm{T}}(\lambda I_n - A_n)^{-1} e_n p(\lambda) \\ &= -\alpha_0\alpha_n[\operatorname{adj}(\lambda I_n - A_n)]_{1n} = (-1)^n\alpha_0\alpha_n\alpha_1\alpha_2\cdots\alpha_{n-1}, \end{aligned}$$

and

$$\begin{aligned} (1-\alpha)p(\lambda) &= p(\lambda) - \alpha_0^2 e_1^{\mathrm{T}}(\lambda I_n - A_n)^{-1} e_1 p(\lambda) + \alpha_n^2 e_n^{\mathrm{T}}(\lambda I_n - A_n)^{-1} e_n p(\lambda) \\ &= p(\lambda) - \alpha_0^2[\operatorname{adj}(\lambda I_n - A_n)]_{11} + \alpha_n^2[\operatorname{adj}(\lambda I_n - A_n)]_{nn} \\ &= p(\lambda) - \alpha_0^2 p_{2n}(\lambda) + \alpha_n^2 p_{1,n-1}(\lambda) \\ &= \left(\lambda^2 + \alpha_0^2\alpha_n^2\right)p_{2,n-1}(\lambda) + \lambda\alpha_{n-1}^2 p_{2,n-2}(\lambda) + \lambda\alpha_1^2 p_{3,n-1}(\lambda) \\ &\quad + \alpha_1^2\alpha_{n-1}^2 p_{3,n-2}(\lambda) \end{aligned}$$

using the definition of $p_{ik}(\lambda)$. Thus $r(\lambda)$ depends on the product $\alpha_0\alpha_n$ but not on the individual values of α_0 and α_n; that is, the eigenvalues of B that are not eigenvalues of A_n depend only on the product $\alpha_0\alpha_n$. Writing $B = B(\alpha_0, \alpha_1, \ldots, \alpha_n)$, for any $t > 0$ define

$$B_t = B(\alpha_0 t, \alpha_1, \ldots, \alpha_{n-1}, \alpha_n/t).$$

Then

$$r(\lambda) = \det(\lambda I_{(k+1)n} - B_t)/\det(\lambda I_n - A_n).$$

Consider the decomposition $B_t = H_t + K_t$ into its skew symmetric part K_t and its symmetric part $H_t = \operatorname{diag}(-\alpha_0^2 t^2, 0, \ldots, 0, \alpha_n^2/t^2)$. If $B_t x = \mu x$ with $x^* x = 1$, then $\mu = x^* H_t x + x^* S_t x$, and well known results on the field of values (see, e.g., [8, p. 140]) give that the eigenvalues of B_t (and B) that are not eigenvalues of A_n have real part in $[-\alpha_0^2 t^2, \alpha_n^2/t^2]$ for all $t > 0$. Since $\bigcap_{t>0}[-\alpha_0^2 t^2, \alpha_n^2/t^2] = \{0\}$, this implies that all of these eigenvalues have real part zero, completing the proof. □

Note that the previous theorem shows that $T_{(k+1)n}$ allows at least nk purely imaginary eigenvalues. This can be applied to prove the following inertia result (which also follows from the numerical results of Section 3).

Corollary 2.4. *The sign pattern T_8 allows any inertia.*

Proof. Let $i(T_8) = (n_1, n_2, n_3)$. In [2, Section 3] matrices A_8 are constructed that give $i(T_8) = (8 - k - n_3, k, n_3)$ for all $n_3 \in \{0, 1, 2, 7, 8\}$. By taking $C = \mathrm{diag}(A_4, A_4)$ and applying the result of Theorem 2.3 with $(k = 1)$, it follows that T_8 allows at least 4 purely imaginary eigenvalues. Moreover, since the parameters of A_4 can be chosen so that A_4 has any spectrum [2, Theorem 9], Theorem 2.3 shows that T_8 allows any inertia with $n_3 \geqslant 4$. The fact that inertias (0, 5, 3), (1, 4, 3), (2, 3, 3) and hence (5, 0, 3), (4, 1, 3) and (3, 2, 3) are possible for T_8 is shown by constructing numerical examples of A_8. $\square$

3. Spectral results for T_n, $8 \leqslant n \leqslant 16$

In [2, Theorem 11] a method is developed to show that T_n allows any spectrum, and this method is applied to prove the result for $2 \leqslant n \leqslant 7$. The first step is to construct a nilpotent $A_n(a_0, \ldots, a_n)$ with $a_n = 1$ by solving n nonlinear equations for $a_j > 0$ arising from the equation $p(\lambda) = \lambda^n$. The second step is to compute the n-by-n Jacobian of this polynomial system and, if this is nonzero, then the Implicit Function Theorem can be used to yield A_n with any desired $p(\lambda) = \sum_{j=0}^{n-1} c_j \lambda^j + \lambda^n$ with sufficiently small coefficients c_j. Thus, given any spectrum, a suitable small positive multiple of that spectrum has such a characteristic polynomial $p(\lambda)$. A rescaling of A_n thus gives a matrix in the pattern T_n having the given spectrum.

We have used MAPLE to extend this procedure to $8 \leqslant n \leqslant 16$, but it could be applied with larger values of n. To find a nilpotent matrix for $n = k$, the parameters a_j for $n = k - 1$ were used to find good starting values. The results for the nilpotent $A_n(a_0, \ldots, a_n)$ matrices are displayed below to 6 decimal places (A_n is symmetric about the reverse diagonal by Theorem 2.1, thus $a_0 = 1$ and only $\lfloor \frac{n}{2} \rfloor$ values need be given).

n	a_1	a_2	a_3	a_4	a_5	a_6	a_7	a_8
8	0.351153	0.082392	0.046671	0.039566				
9	0.347296	0.078725	0.041889	0.032088				
10	0.344576	0.076231	0.038841	0.027807	0.025085			
11	0.342584	0.074450	0.036764	0.025091	0.021108			
12	0.341081	0.073132	0.035276	0.023244	0.018599	0.017332		
13	0.339918	0.072126	0.034169	0.021921	0.016899	0.014964		
14	0.338999	0.071341	0.033322	0.020938	0.015685	0.013365	0.012695	
15	0.338261	0.070716	0.032658	0.020183	0.014783	0.012227	0.011170	
16	0.337658	0.070209	0.032126	0.019591	0.014092	0.011383	0.010088	0.009700

Note that each coefficient a_j appears to be converging to a fixed value as n increases, and that $a_{j+1} < a_j$ for $1 \leqslant j \leqslant \lfloor \frac{n}{2} \rfloor$. The n-by-n Jacobian is evaluated numerically at these nilpotent values. For $8 \leqslant n \leqslant 16$, all Jacobians are nonzero; thus we have the following result.

Theorem 3.1. *For* $2 \leqslant n \leqslant 16$, *the sign pattern* T_n *allows any spectrum.*

4. Low rank perturbations of general matrices

The result of Theorem 2.3 and the fact that every eigenvalue of A_n has geometric multiplicity 1, lead us to consider low rank perturbations of more general matrices. In the following a void sum is understood to be 0.

Theorem 4.1. *Let* $X \in \mathbb{C}^{n,n}$ *have distinct eigenvalues* $\lambda_0, \ldots, \lambda_w$. *Let* $m_{1i} \geqslant m_{2i} \geqslant \cdots \geqslant m_{ni} \geqslant 0$ *be the dimensions of the Jordan blocks associated with the eigenvalue* λ_i, $i = 0, \ldots, w$ (*note that* $m_{1i} > 0$ *but other* m_{vi} *may be zero*). *Denote the algebraic multiplicity of* λ_i *by*

$$a_i = m_{1i} + \cdots + m_{ni} = \sum_{v=1}^{n} m_{vi}.$$

If $\operatorname{rank}(U) = p$, *then* λ_i *is an eigenvalue of* $X + U$ *with algebraic multiplicity* $\tilde{a}_i$, *where*

$$\sum_{v=p+1}^{n} m_{vi} = a_i - \sum_{v=1}^{p} m_{vi} \leqslant \tilde{a}_i \leqslant a_i + \sum_{\substack{j=0\\ j\neq i}}^{w} \sum_{v=1}^{p} m_{vj}.$$

Furthermore, each bound can be attained for a suitable perturbation U *with* $\operatorname{rank}(U) = p$.

Proof. For given λ_i, assume that

$$m_{1i} \geqslant m_{2i} \geqslant \cdots \geqslant m_{ri} > 0, \quad m_{r+1,i} = 0.$$

We first prove the left inequality, and to simplify notation the subscript i is dropped. Wlog $\lambda = 0$ by a scalar shift. Define for $s = 1, 2, \ldots,$

$$n_s = \dim N(X^s), \quad \tilde{n}_s = \dim N((X+U)^s),$$

where $N(Y)$ is the null space of Y. Then from the Jordan normal form (see, for example, [6, p. 131])

$$a = \max_s n_s, \quad \tilde{a} = \max_s \tilde{n}_s$$

with $n_1 < n_2 < \cdots < n_{m_1} = n_{m_1+1}$, giving $a = n_{m_1}$. From the well known relations $\operatorname{rank}(A+B) \leqslant \operatorname{rank}(A) + \operatorname{rank}(B)$, $\operatorname{rank}(AB) \leqslant \min\{\operatorname{rank}(A), \operatorname{rank}(B)\}$, and

$$(X+U)^t - X^t = \sum_{v=0}^{t-1} (X+U)^v U X^{t-v-1}$$

for $t = 1, 2, \ldots,$ it follows that $\operatorname{rank}((X+U)^t - X^t) \leqslant tp$. Thus $|\tilde{n}_t - n_t| \leqslant tp$, since $|\tilde{n}_t - n_t| = |\operatorname{rank}(X+U)^t - \operatorname{rank}(X^t)| \leqslant \operatorname{rank}((X+U)^t - X^t)$. Hence for any t,

$$\tilde{a} \geqslant \tilde{n}_t \geqslant n_t - tp = a - (a - n_t) - tp. \tag{4}$$

We now obtain the best possible lower bound for $\tilde{a}$ by minimizing the expression $tp + (a - n_t)$, and we claim that this minimum is achieved when $t = m_{p+1}$. To see this, first note that if $1 \leqslant \nu \leqslant r - 1$, then

$$n_{\tau+1} = n_\tau + \nu \tag{5}$$

for all τ such that $m_{\nu+1} \leqslant \tau < m_\nu$. This follows since for the specified values of τ, there are exactly ν Jordan blocks of length $> m_{\nu+1}$ that contribute to an increase in the magnitude of the values $n_{\tau+1}$. Now from (5) by summing over all values of τ on the given interval

$$n_{m_\nu} - n_{m_{\nu+1}} = \nu(m_\nu - m_{\nu+1}) \quad \text{for } 1 \leqslant \nu \leqslant r - 1.$$

Taking $t = m_{p+1}$ gives

$$\begin{aligned} & m_{p+1}\, p + (a - n_{m_{p+1}}) \\ & \quad = m_{p+1} p + (n_{m_1} - n_{m_2}) + (n_{m_2} - n_{m_3}) + \cdots + (n_{m_p} - n_{m_{p+1}}) \\ & \quad = m_{p+1} p + (m_1 - m_2) + 2(m_2 - m_3) + \cdots + p(m_p - m_{p+1}) \\ & \quad = m_1 + m_2 + \cdots + m_p. \end{aligned}$$

Using this equality in (4) gives

$$\tilde{a} \geqslant \sum_{v=p+1}^{r} m_v = \sum_{v=p+1}^{n} m_v,$$

proving the left inequality. To see that the above inequality is the best possible, consider a specific perturbation U. Assume that $X = \operatorname{diag}(J_{m_1}, \ldots, J_{m_r}, K)$ where J_{m_v}, $v = 1, \ldots, r$, are the Jordan blocks of dimensions m_v associated with $\lambda = 0$, and the (possibly void) block K is nonsingular. Then putting a 1 in position $(m_v, 1)$ of the block J_{m_v}, $v = 1, \ldots, p$, makes these blocks nonsingular, so the resulting matrix has an eigenvalue $\lambda = 0$ with algebraic multiplicity $\sum_{v=p+1}^{r} m_v$.

For the right inequality, applying the above result to λ_j gives

$$a_j - \tilde{a}_j \leqslant \sum_{v=1}^{p} m_{vj}.$$

The inequality $\sum_{j=0}^{w} \tilde{a}_j \leqslant n = \sum_{j=0}^{w} a_j$ gives

$$\begin{aligned} \tilde{a}_i &\leqslant \sum_{j=0}^{w} a_j - \sum_{\substack{j=0 \\ j \neq i}}^{w} \tilde{a}_j = a_i + \sum_{\substack{j=0 \\ j \neq i}}^{w} (a_j - \tilde{a}_j) \\ &\leqslant a_i + \sum_{\substack{j=0 \\ j \neq i}}^{w} \sum_{v=1}^{p} m_{vj}, \end{aligned}$$

completing the proof of the right inequality. To consider equality of the right side, wlog take $\lambda_0 = 0$. Matrix X is similar to $\operatorname{diag}(\overline{J}_0, \overline{J}_1, \ldots, \overline{J}_w)$ where $\overline{J}_j$ is the Jordan normal form associated with $\lambda_j,\ j = 0, \ldots, w$. Each $\overline{J}_j$ for $j = 1, \ldots, w$ has

Jordan blocks $J_{m_{1j}}, \ldots, J_{m_{nj}}$ where $\overline{J}_j = \mathrm{diag}(J_{m_{1j}}, \ldots, J_{m_{nj}})$ and $J_{m_{qj}}$ is vacuous if $m_{qj} = 0$. For $v = 1, \ldots, p$, consider the direct sum of the vth largest Jordan block of each eigenvalue, namely $\mathrm{diag}(J_{m_{v1}}, \ldots, J_{m_{vw}})$, which is similar to a companion matrix [6, p. 155]. Each companion matrix can be made nilpotent by adding a rank 1 matrix. Hence by adding a rank p matrix U, the algebraic multiplicity of $\lambda_0 = 0$ in $X + U$ can be increased by $\sum_{j=1}^{w} \sum_{v=1}^{p} m_{vj}$, giving the equality result. $\square$

Note that if λ_i is not an eigenvalue of X, then the upper bound in Theorem 4.1 may still give a nontrivial result for the algebraic multiplicity of λ_i in the perturbed matrix $X + U$. The lower bound in Theorem 4.1 can be found from [9, Theorem 4]. For $\lambda = 0$, the sequence $m_1, m_2, \ldots, m_r$ is called the Segré characteristic of X, and the vector of the differences $n_{s+1} - n_s$ is called the height characteristic of X; see, for example, [5].

In the following result, Theorem 4.1 is applied to a block diagonal matrix to obtain bounds on the algebraic multiplicity of an eigenvalue. This is useful if the Jordan structure of the block diagonal matrices is unknown.

Corollary 4.2. *Let $X = \mathrm{diag}(X_1, \ldots, X_q) \in \mathbb{C}^{n,n}$ be a block diagonal matrix and suppose that λ_i is an eigenvalue of X_v with algebraic multiplicity $\tilde{m}_{vi}$ for $1 \leqslant v \leqslant q$ where $\tilde{m}_{1i} \geqslant \cdots \geqslant \tilde{m}_{qi}$. Denote the algebraic multiplicity of λ_i in X by $a_i = \sum_{i=1}^{q} \tilde{m}_{vi}$. If $\mathrm{rank}(U) = p$, then λ_i is an eigenvalue of $X + U$ with algebraic multiplicity $\tilde{a}_i$, where*

$$\sum_{v=p+1}^{n} \tilde{m}_{vi} \leqslant \tilde{a}_i \leqslant a_i + \sum_{\substack{j=0 \\ j \neq i}}^{w} \sum_{v=1}^{p} \tilde{m}_{vj}.$$

Proof. The Jordan structure of X is the union of the Jordan structures of the X_v. Using the same notation for the Jordan structure of X as in the statement of Theorem 4.1,

$$\sum_{v=1}^{p} m_{vi} \leqslant \sum_{v=1}^{p} \tilde{m}_{vi} \quad \text{for } p = 1, 2, \ldots, \max\{r, q\}$$

(defining $m_{vi} = 0$ for $v > r$, and $\tilde{m}_{vi} = 0$ for $v > q$). In addition

$$a_i = \sum_{v=1}^{n} m_{vi} = \sum_{v=1}^{n} \tilde{m}_{vi},$$

and the result follows from Theorem 4.1. $\square$

Note that the case considered in Theorem 2.3 serves as an example for the case of equality of the lower bound in Corollary 4.2 (with $p = k$, $q = k + 1$, $\tilde{m}_{v1} = \cdots = \tilde{m}_{vq}$). However, because of the special structure of A_n we are able to prove more about the spectrum of the perturbed matrix in the case of Theorem 2.3. The lower bound of Corollary 4.2 with $q = 2$ was used to construct the numerical examples of A_8 with inertias $(0, 5, 3)$, $(1, 4, 3)$ and $(2, 3, 3)$ reported in Corollary 2.4.

Acknowledgements

The authors thank David Harrington for advice about the MAPLE calculations in Section 3. L. Elsner thanks the University of Victoria, where this work was carried out during his sabbatical leave.

References

[1] M.T. Chu, Inverse eigenvalue problems, SIAM Rev. 40 (1998) 1–39.
[2] J.H. Drew, C.R. Johnson, D.D. Olesky, P. van den Driessche, Spectrally arbitrary patterns, Linear Algebra Appl. 308 (2000) 121–137.
[3] Y. Gao, Y. Shao, Inertially arbitrary patterns, Linear and Multilinear Algebra 49 (2001) 161–168.
[4] L. Elsner, D. Hershkowitz, On the spectra of close-to-Schwarz matrices, Linear Algebra Appl. 363 (2003) 81–88.
[5] D. Hershkowitz, The combinatorial structure of generalized eigenspaces—from nonnegative matrices to general matrices, Linear Algebra Appl. 302–303 (1999) 173–191.
[6] R.A. Horn, C.R. Johnson, Matrix Analysis, Cambridge University Press, 1985.
[7] J.J. McDonald, D.D. Olesky, M.J. Tsatsomeros, P. van den Driessche, On the spectra of striped sign patterns, Linear and Multilinear Algebra 51 (2003) 39–48.
[8] M. Marcus, H. Minc, A Survey of Matrix Theory and Matrix Inequalities, Allyn and Bacon, Boston, 1964.
[9] J. Moro, F.M. Dopico, First order eigenvalue perturbation theory and the Newton diagram, in: Z. Drmač et al. (Eds.), Proceedings of the Second Conference on Applied Mathematics and Scientific Computing, Dubrovnik, Croatia, Applied Mathematics and Scientific Computing, Kluwer Academic Publishers, 2002, pp. 143–175.